ATTACHMENT A

RESPONSES TO EPA LETTER 8HWM-FF - STATISTICAL COMPARISONS TO BACKGROUND AT ROCKY FLATS DATED SEPTEMBER 21, 1993 AND TO CDH LETTER - STATISTICAL METHODS FOR THE COMPARISON OF REMEDIAL INVESTIGATION DATA TO BACKGROUND DATA AT ROCKY FLATS PLANT, DATED SEPTEMBER 13, 1991

RESPONSES TO EPA LETTER 8HWM-FF - STATISTICAL COMPARISONS TO BACKGROUND AT ROCKY FLATS DATED SEPTEMBER 21, 1993:

SPECIFIC COMMENTS

1. Page 2, Seventh Bullet. It is suggested that the same field sampling and laboratory procedures be used for both background and site data. The statement should be extended to include data aggregation. Past review of RFP data from operable units showed inconsistencies in the methodology used to aggregate data. Problems encountered at this phase will be magnified at later stages of the background analysis

Clarification. Data aggregation is another topic, being addressed by CDH and EPA separately from this forum, which deals strictly with site-to-background comparison.

2 Page 4, Task 1, Observation 1, Third Bullet. This statement suggests that background analysis should be the initial state in selecting COCs. This is consistent with the COC selection methodology developed for Rocky Flats by DOE, EPA, and CDH. However, in order to manage DOE's effort in background comparisons, we point out that it is not necessary to carry all chemicals through an elaborate, time consuming statistical analysis if they can be eliminated as essential nutrients or as infrequently detected chemicals It may be more cost-effective and expeditious to simply eliminate chemicals on the basis of these two preliminary criteria than to conduct a background analysis only to eliminate them later based on the background analysis We suggest that DOE consider this in the development of a plan to implement Dr. Gilbert's approach

Concur. CDH is correct that time might be saved in eliminating nutrients and infrequently detected analytes prior to statistical analysis. We will investigate whether significant time is saved by following CDH's recommendation, and if so, will adopt the suggestion.

3 Page 5, Task 1, Observation 4, Second Bullet. This statement expresses concern about measurements that are less than the contract required detection limits (CRQL) but above instrument detection limits (IDL) According to Risk Assessment Guidance for Superfund, Human Health Evaluation Manual, Volume I, Part A, these measurements should be "J" coded and interpreted as estimated values. They should not be viewed as non-detected chemicals If they are currently classified as non-detect chemicals in the RFP background geochemical report, the entire validation process currently in place should be reevaluated.

Clarification. There has been confusion over the detection limits and their application A qualifier of "J" indicates that the reported value is between the instrument detection limits and the contract required detection limits A non-detect has a reported value of a detection limit, not the detected value, and conveys less information than a "J"

4 Page 9, Paragraphs 3 and 4. The essence of this discussion is that a hot measurement (HM) concentration should serve as a "safety net" that can prevent "hot spots" from passing

unnoticed in a risk assessment. It should be noted that this need has been previously recognized and was addressed in the original flow chart devised during the summer 1992 meetings involving EPA, DOE, and CDH. At that time, it was agreed that a risk-based concentration (RBC) would effectively serve as the "hot measurement " Although a UTL has some utility in identifying hot spots, there is no need to conduct a lengthy analysis if the highest detected concentrations do not exceed a predetermined RBC and pose an unacceptable human health risks. Thus, it is possible to have measurements above the UTL but below an RBC in which case there would be little reason to consider the chemical further.

Clarification. The Guide for Conducting Statistical Comparisons of RFI/RI Data and Background Data at the Rocky Flats Plant (called The Guide subsequently) addresses statistical determination of the presence or absence of analytes, and does not address human health effects. For each OU, additional tests will determine if the analyte concentrations present are below regulatory (ARARs) and/or human health effect (PRGs) levels, but that is external to the statistical discussion at hand.

5 Page 10, Third and Fourth Bullet. This statement refers to lowering the potential for a Type I, false positive error to using a 99 percent UTL on the 99 percentile. However, this concern is not properly balanced against the potential for a Type II error A false negative could have profound consequences on the risk assessment and subsequent remedy selected for the site.

Do not concur. If the 95% UTL were used, then a very high percentage of data points would be considered pCoCs, because theoretically, even a background population will have 5% of readings above the UTL. A site, even if its concentration levels are slightly above background, may have considerably more than 5% of its readings above the $UTL_{95/95}$ Any analytes that show a false negative on this test will still be considered pCoCs if they test positive on any of the other statistical tests.

6 Page 11, Second Paragraph. This paragraph suggests that data quality objectives (DQOs) be established at the design stage of the studies Although this is a relevant comment in the context of planning a background analysis, the background and most of the OU planning and sampling has already been completed. Thus, this comment is appropriate in theory but there is little chance for implementation Revitalized effort should be directed to establishing DQOs where they were not previously established, and analyzing whether the sampling efforts completed to date have succeeded in meeting these DQOs DOE, EPA, and CDH will need to look at options for correcting the situation if the DQOs have not been met

Concur. The draft RIs for each OU have a section for reviewing data quality Each OU manager bears the responsibility for ensuring that DQOs are met for his or her OU.

7. Task 4, Flow Chart for Comparing OU Data to Background. With a minor exception, this flow chart adequately describes the framework for a background analysis The exception is an inadequate description of appropriate conditions under which particular statistical tests should applied

Explicit guidelines for the application of specific statistical tests under well-defined conditions

should be presented to circumvent future misunderstandings· It would be highly useful for EPA, DOE, and CDH to agree to a predetermined paradigm in which all possible circumstances and conditions have been anticipated and the appropriate statistical tests identified. Knowing in advance what particular test will be applied under what circumstances will prevent protracted discussions and possible disagreements.

Concur. The Background Comparison Methodology chart shows the specific tests and gives the conditions under which they are or are not applicable. In addition, The Guide's text states which tests will be conducted, under what circumstances

IMPLEMENTATION ISSUES

1 EPA, DOE, and CDH must reach consensus on procedures for defining non-detects

Concur. The Guide states that non-detects will be considered to be one-half of the detection limit, in accordance with EPA guidance.

2. EPA, DOE, and CDH must reach consensus on what hot measurement value should be used

Concur. Our methodology uses a value of $UTL_{99/99}$

3. EPA, DOE, and CDH must establish data quality objectives which address acceptable power and confidence levels, required detection limits, and anticipated data aggregation.

Concur. The draft RIs for each OU have a section for reviewing data quality. Each OU manager bears the responsibility for ensuring that DQOs are met for his or her OU.

4. EPA, DOE, and CDH must revisit the assumptions which Dr. Gilbert lists on page two of his cover letter. Are these assumptions valid? What are the consequences if the assumptions are violated? Can this be handled in an uncertainty analysis?

Clarification. All of the assumptions listed, except for the last four, are difficult to quantify and are thus not "valid" or "invalid" These last four are now answered individually.

The same field-sampling techniques are used for background and site, so this assumption is valid

Measurements are not always validated by subcontractors before the draft RFI/RI statistical testing has been completed, so this assumption is not valid. When the data validation results have been obtained, the data are reanalyzed, and the final RFI/RI contains no invalidated data

Background data were checked for outliers, per EPA comments upon the 1992 Background Geochemical Report, and extreme outliers were excluded from statistical analysis in the 1993 Backgroun Geochemical Report, so this assumption is not entirely valid However, OU data outliers are not typically deleted, although data from the OUs are checked for "geochemical reasonableness", and any unusual results are discussed in the ensuing reports

The instrument detection limits are not always reported in the data bases, so this assumption is not completely valid However, the costs of recovering this information would be considerable

5. EPA, DOE, and CDH must reach consensus on a paradigm for implementation The issues to be worked out include:

a The appropriate background data sets by analyte, medium, and location

Concur. The section of The Guide entitled "Determine Background and OU Target Populations" addresses how this will be done.

b. How to deal with clearly non-random (e.g., spatial) patterns.

Concur. The Guide states in the Professional Judgement section that spatial patterns are subject to professional judgement, which is then subject to EPA and CDH review.

c. Measurement errors and multiple non-detects.

Concur. Measurement errors are an inevitable part of physical data. Efforts are taken throughout the data-collection process to minimize errors. Multiple non-detects are dealt with by replacing the data value with ½ of the reported value, or by using the Gehan test.

d. Structure for the formal statistical tests.

Concur. The Guide furnishes this structure

e. Data aggregation for comparison in the statistical tests

Clarification. Data aggregation is another topic, being addressed by CDH and EPA separately from this forum, which deals strictly with site-to-background comparison